

## Scatterplots

### Notes:

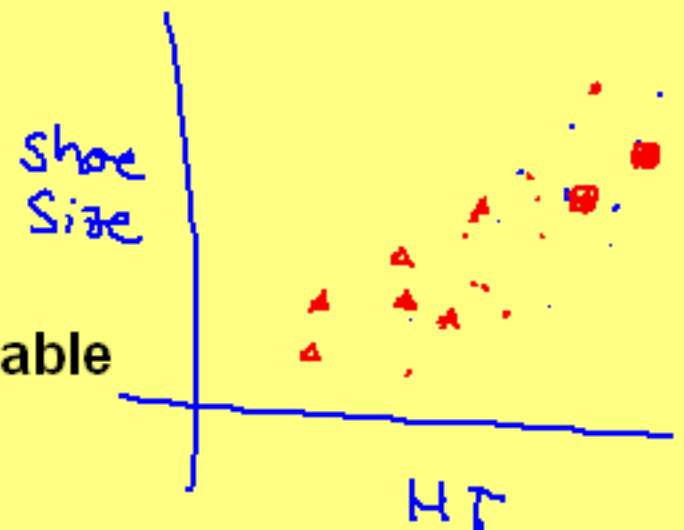
- Shows the relationship between 2 quantitative variables
- Can show categorical variables by (1) colors/shapes
- Individuals are represented by the dots on the plot  
*#duplicates? hard to see*

- Explanatory Variable:

- On the X axis *the researcher determines/changes*
- Explains or causes the change in the y variable

- Response Variable:

- On the y axis
- Measures the outcome of an experiment or study



Ex: medication (+)  
# tumors (y)

# Interpreting Scatterplots:

## Overall pattern

- Get a sense of what the data/plot looks like in general, then comment on the following 3 things

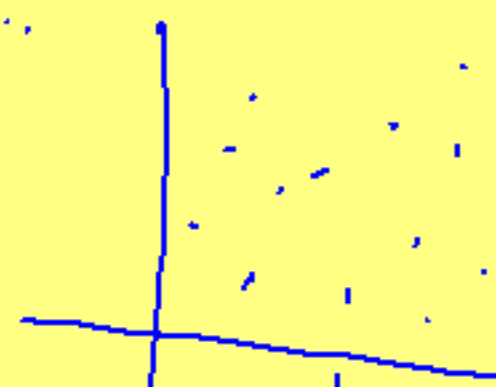
(1) Form  
\* Linear:



Curved:



Scattered:



(2) Direction \* Overall  
Negative Association

as  $x \uparrow$ ,  $y \downarrow$



Positive Association

as  $x \uparrow$ ,  $y \uparrow$



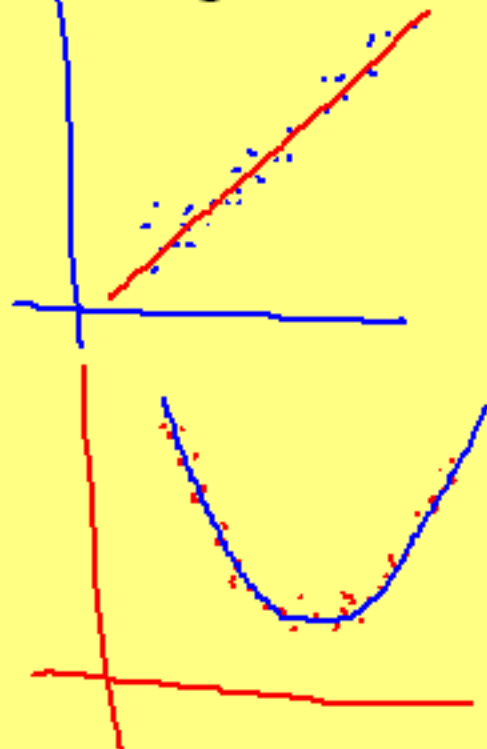
(1) Strength-

- how closely the pts follow a simple form.
- Use the following words (or combinations of these):
  - strong
  - moderate
  - weak
  - scattered

moderately strong

Examples:

Strong



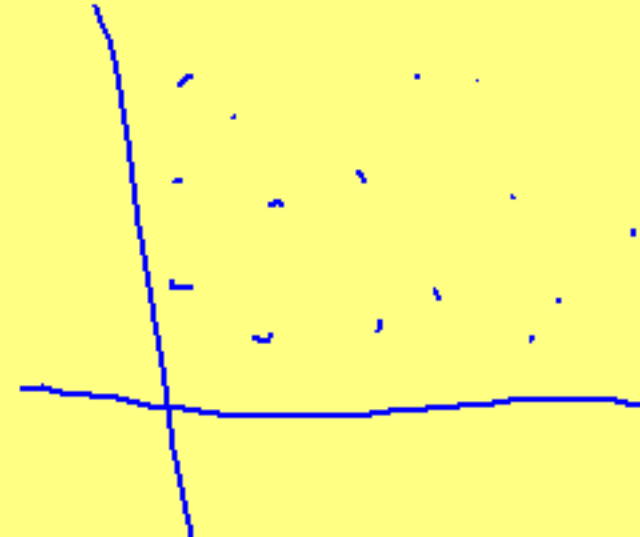
Moderate



Weak



Scattered



## Examples: Describing Scatterplots

1) 2) 3)

\* You try the next 4 examples

\* Now complete worksheet 2.1

① - curved / parabolic / quadratic

- neg. assoc. until  $x = \underline{\hspace{1cm}}$   
then pos. assoc.

- strong

② - linear / curved

- neg.

- mod. ~~weak~~

③ - linear

- pos.

- weak



④ - 2 clusters of data

- scattered

- negative

- weak

- poss. outlier

## Worksheet 2.1- ANSWERS

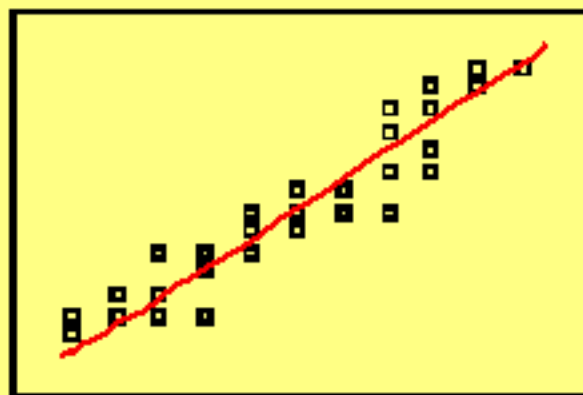
① Chart:

|   | Strong | Mod | Weak |
|---|--------|-----|------|
| - | C      | D   | F    |
| + | E      | A   | B    |

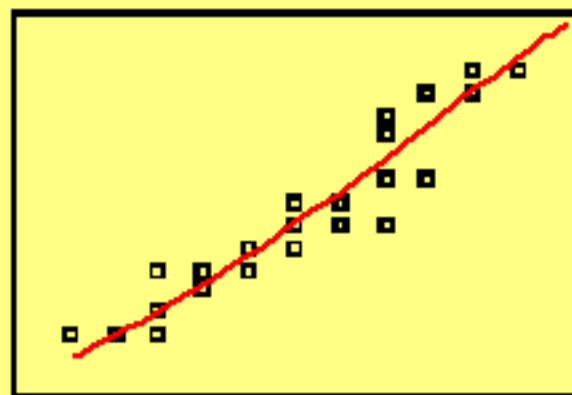
- ②
- |                   |                   |
|-------------------|-------------------|
| a) +, strong      | f) -, mod.        |
| b) +, mod. strong | g) +, strong      |
| c) Scattered      | h) +, mod. strong |
| d) +, mod.        |                   |
| e) -, strong      |                   |

WHAT PLOT HAS THE STRONGER RELATIONSHIP? BY HOW MUCH??

A



B



## Correlation:

Symbol:

$r$

Definition:

the strength & direction of  
the linear relationship  
btw. 2 variables.

Formula:

$$r = \frac{1}{n-1} \sum \left( \frac{x_i - \bar{x}}{S_x} \right) \left( \frac{y_i - \bar{y}}{S_y} \right)$$

- Grouping/ Ungrouping
- How to make scatterplots on the calculator
- Deleting a point from a list
- Using the program CORR

std.  
dev.

← standardizing  
\* took away units

$$Z = \frac{X - \mu}{\sigma}$$
$$Z = 1.4$$



**CORRELATION COEFFICIENT: the last few notes....**

## COEFFICIENT OF DETERMINATION:

What is it?

How do we interpret it?